```
Set     Items    Description
S1      587114   RECORD? ? OR DOCUMENT? OR PUBLICATION? ? OR PAGE? OR DATAF-
                 ILE? OR FLATFILE OR FLAT()FILE?
S2      2365318  FIELD? OR AREA? OR REGION? OR SEGMENT? OR PHRASE? OR SENTE-
                 NC? OR PARAGRAPH?
S3      3652295  COMPAR? OR MATCH? OR FIND? OR LOCAT? OR SAME? OR IDENTICAL?
                 OR EQUIVALENT?
S4      15723    S1(2N)(TWO OR SECOND OR ANOTHER OR PAIR OR 2ND OR PLURAL?)
S5      1300837  COMBIN? OR MERG? OR INTEGRAT?
S6      426778   TEMPORAR? OR BUFFER? OR CACHE?
S7      3499967  OUTPUT? OR SEND? OR DELIVER? OR DOWNLOAD? OR PRINT?
S8      589519   PARTIAL? OR SEGMENT? OR MATCHING()(AREA? OR PHRASE? OR TER-
                 M? OR FIELD?)
S9      1711185  DELET? OR REMOV? OR WIPE? OR ERAS? OR SCRUB?
S10     12       S4 AND S2 AND S3 AND S5 AND S6
S11     2        S4 AND S2 AND S3 AND S7 AND S8 AND S9
S12     6        S3 AND S4 AND S7 AND S8 AND S9
S13     400      S3 AND S4 AND S5
S14     35       S8 AND S13
S15     45       S10 OR S11 OR S12 OR S14
S16     38       S15 NOT AD=20010111:20030111
S17     36       S16 NOT AD=20030111:20050701
S18     36       IDPAT (sorted in duplicate/non-duplicate order)
S19     36       IDPAT (primary/non-duplicate records only)
S20     116      S13 AND S2
S21     12       S20 AND IC=G06F-007
S22     9        S21 NOT S15
S23     7        S22 NOT AD>20010117
S24     1336702  TEXT? OR DOCUMENT? OR WORD? ? OR PHRASE? OR PARAGRAPH? OR -
                 TERMS OR TERM OR CHARACTER()STRING? OR SEGMENT?
S25     20127    S3(2N)S24
S26     121      S5 AND S6 AND S25
S27     67       S26 AND IC=G06F
S28     11243    S24(5N)S5
S29     21       S27 AND S28
S30     20       S29 NOT S15
S31     19       S30 NOT AD=20010111:20030111
S32     18       S31 NOT AD=20030111:20050701
S33     18       IDPAT (sorted in duplicate/non-duplicate order)
S34     18       IDPAT (primary/non-duplicate records only)
File 347:JAPIO Nov 1976-2005/Feb(Updated 050606)
         (c) 2005 JPO & JAPIO
File 350:Derwent WPIX 1963-2005/UD,UM &UP=200537
         (c) 2005  Thomson Derwent
```

05429684   **Image available**
KANA-KANJI CONVERTER

PUB. NO.:   09-044484 [JP 9044484 A]
PUBLISHED: February 14, 1997 (19970214)
INVENTOR(s):   OIKE YOKO
APPLICANT(s):   BROTHER IND LTD [000526] (A Japanese Company or Corporation),
                JP (Japan)
APPL. NO.:   07-192973 [JP 95192973]
FILED: July 28, 1995 (19950728)
INTL CLASS: [6] G06F-017/22
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications)
JAPIO KEYWORD:R139 (INFORMATION PROCESSING -- Word Processors)

ABSTRACT

PROBLEM TO BE SOLVED: To improve conversion efficiency by rewriting the
contents of a pertinent conversion result storage means in accordance with
the rewriting information on a retrieved rule.

SOLUTION: A retrieval processing as to whether the rule having the
 combination of the **word** patterns **matching** with the **combination** of
the **words** of a conversion result storage area 31 exists in a rule
dictionary 49 or not is performed by the rule retrieval program 143 in a
ROM 40. When the required rule exists, further, whether the rule selection
flag of the rule is active or nonactive is checked by a rule selection
program 45. When the flag is active, the **combination** part of the **words**
 matching with the **combination** of the **word** patterns of the rule of the
words in the conversion result storage area 31 is rewritten by the rule
rewriting program 44 in the ROM 40 in accordance with the rewriting
information on the rule. Thus, a rule conversion processing is terminated,
the contents of the conversion result storage area 31 is stored in an
output **buffer** area 33 and the candidate display of the contents on an
output device 50 is performed.

```
Set      Items    Description
S1     1819328    RECORD? ? OR DOCUMENT? OR PUBLICATION? ? OR PAGE? OR DATAF-
                  ILE? OR FLATFILE OR FLAT()FILE?
S2     9836983    FIELD? OR AREA? OR REGION? OR SEGMENT? OR PHRASE? OR SENTE-
                  NC? OR PARAGRAPH?
S3    12450681    COMPAR? OR MATCH? OR FIND? OR LOCAT? OR SAME? OR IDENTICAL?
                  OR EQUIVALENT?
S4      26902     S1(2N)(TWO OR SECOND OR ANOTHER OR PAIR OR 2ND OR PLURAL?)
S5     4979493    COMBIN? OR MERG? OR INTEGRAT?
S6      440004    TEMPORAR? OR BUFFER? OR CACHE?
S7     1993424    OUTPUT? OR SEND? OR DELIVER? OR DOWNLOAD? OR PRINT?
S8     1688303    PARTIAL? OR SEGMENT? OR MATCHING()(AREA? OR PHRASE? OR TER-
                  M? OR FIELD?)
S9     1246216    DELET? OR REMOV? OR WIPE? OR ERAS? OR SCRUB?
S10          1    S4 AND S2 AND S3 AND S5 AND S6
S11          2    S4 AND S2 AND S3 AND S7 AND S8 AND S9
S12          3    S3 AND S4 AND S7 AND S8 AND S9
S13        982    S3 AND S4 AND S5
S14         73    S8 AND S13
S15         76    S10 OR S11 OR S12 OR S14
S16    3894081    TEXT OR TEXTUAL OR DOCUMENT? OR WORD? ? OR PHRASE? OR PARA-
                  GRAPH? OR TERM OR TERMS OR (CHARACTER OR TEXT)(N)(STRING? OR -
                  SEGMENT?)
S17      68276    S3(2N)S16
S18      26648    S16(2N)(MATCHING OR SIMILAR OR SAME OR IDENTICAL OR EQUAL -
                  OR EQUIVALENT)
S19          1    S18 AND S4 AND S5 AND (S6 OR WORKSPACE? OR WORK()(SPACE? OR
                  AREA?) OR CLIPBOARD?)
S20          1    S18 AND S4 AND (S6 OR WORKSPACE? OR WORK()(SPACE? OR AREA?)
                  OR CLIPBOARD?)
S21         47    S18 AND S5 AND (S6 OR WORKSPACE? OR WORK()(SPACE? OR AREA?)
                  OR CLIPBOARD?)
S22        291    S18  AND (S6 OR WORKSPACE? OR WORK()(SPACE? OR AREA?) OR C-
                  LIPBOARD?)
S23         59    S22 AND (S7 OR S8 OR S9)
S24        171    S15 OR S19 OR S20 OR S21 OR S23
S25        132    RD (unique items)
S26         88    S25 NOT PY>2001
File     8:Ei Compendex(R) 1970-2005/Jun W1
           (c) 2005 Elsevier Eng.  Info. Inc.
File    35:Dissertation Abs Online 1861-2005/May
           (c) 2005 ProQuest Info&Learning
File    65:Inside Conferences 1993-2005/Jun W2
           (c) 2005 BLDSC all rts. reserv.
File     2:INSPEC 1969-2005/Jun W1
           (c) 2005 Institution of Electrical Engineers
File    94:JICST-EPlus 1985-2005/Apr W4
           (c)2005 Japan Science and Tech Corp(JST)
File   111:TGG Natl.Newspaper Index(SM) 1979-2005/Jun 14
           (c) 2005 The Gale Group
File     6:NTIS 1964-2005/Jun W1
           (c) 2005 NTIS, Intl Cpyrght All Rights Res
File   144:Pascal 1973-2005/Jun W1
           (c) 2005 INIST/CNRS
File    34:SciSearch(R) Cited Ref Sci 1990-2005/Jun W1
           (c) 2005 Inst for Sci Info
File    99:Wilson Appl. Sci & Tech Abs 1983-2005/May
           (c) 2005 The HW Wilson Co.
File    95:TEME-Technology & Management 1989-2005/May W2
           (c) 2005 FIZ TECHNIK
```

26/5/6   (Item 6 from file: 8)
DIALOG(R)File   8:Ei Compendex(R)

04333973   E.I. No: EIP96013008683
  Title: Evaluating spatial correspondence of zones in document recognition systems
  Author: Garris, Michael D.
  Corporate Source: Natl Inst of Standards and Technology, Gaithersburg, MD, USA
  Conference Title: Proceedings of the 1995 IEEE International Conference on Image Processing. Part 3 (of 3)
  Conference   Location: Washington,   DC,   USA   Conference   Date: 19951023-19951026
  Sponsor: IEEE
  E.I. Conference No.: 44184
  Source: IEEE International Conference on Image Processing v 3 1995. IEEE, Los Alamitos, CA, USA,95CB35819. p 304-307
  Publication Year: 1995
  CODEN: 85QTAW
  Language: English
  Document Type: CA; (Conference Article)   Treatment: T; (Theoretical)
  Journal Announcement: 9603W3
  Abstract: This paper introduces scoring methods developed to automatically assess the performance of document recognition systems; specifically, to evaluate the spatial correspondence of zones produced by a **document segmentor** . **Two** different approaches are discussed. The first approach (based on zone overlap and nearest-neighbors) is better applied to **merged** zones, whereas the second approach (based on zone alignments) is better applied to nested zones (such as those found in tables and graphs). Definitions of coverage and efficiency error are presented, and scoring results on real system output is provided that validates the usefulness of these methods to **compare** different document recognition algorithms. Currently, no standard testing procedures exist for measuring and **comparing** algorithms within a complex document recognition system. Scoring methods, like the ones introduced in this paper, serve as design and validations tools, expediting the development and deployment of document analysis technology for system developers and end users. (Author abstract) 13 Refs.
  Descriptors: *Pattern recognition systems; Imaging systems; Database systems; Image **segmentation** ; Real time systems; Algorithms; Image analysis; Errors; Information technology; Systems analysis
  Identifiers: Document recognition systems; Document **segmentor** ; Reference zones; Scoring methods; Document analysis technology; Pixels; Polygons
  Classification Codes:
  723.5 (Computer Applications); 723.2 (Data Processing); 723.3 (Database Systems); 902.2 (Codes & Standards); 921.4 (Combinatorial Mathematics, Includes Graph Theory, Set Theory)
  723 (Computer Software); 741 (Optics & Optical Devices); 902 (Engineering Graphics & Standards); 921 (Applied Mathematics)
  72 (COMPUTERS & DATA PROCESSING); 74 (OPTICAL TECHNOLOGY); 90 (GENERAL ENGINEERING); 92 (ENGINEERING MATHEMATICS)

```
Set     Items     Description
S1     1819328    RECORD? ? OR DOCUMENT? OR PUBLICATION? ? OR PAGE? OR DATAF-
                  ILE? OR FLATFILE OR FLAT()FILE?
S2     9836983    FIELD? OR AREA? OR REGION? OR SEGMENT? OR PHRASE? OR SENTE-
                  NC? OR PARAGRAPH?
S3    12450681    COMPAR? OR MATCH? OR FIND? OR LOCAT? OR SAME? OR IDENTICAL?
                  OR EQUIVALENT?
S4      26902    S1(2N)(TWO OR SECOND OR ANOTHER OR PAIR OR 2ND OR PLURAL?)
S5     4979493    COMBIN? OR MERG? OR INTEGRAT?
S6      440004    TEMPORAR? OR BUFFER? OR CACHE?
S7     1993424    OUTPUT? OR SEND? OR DELIVER? OR DOWNLOAD? OR PRINT?
S8     1688303    PARTIAL? OR SEGMENT? OR MATCHING()(AREA? OR PHRASE? OR TER-
                  M? OR FIELD?)
S9     1246216    DELET? OR REMOV? OR WIPE? OR ERAS? OR SCRUB?
S10          1    S4 AND S2 AND S3 AND S5 AND S6
S11          2    S4 AND S2 AND S3 AND S7 AND S8 AND S9
S12          3    S3 AND S4 AND S7 AND S8 AND S9
S13        982    S3 AND S4 AND S5
S14         73    S8 AND S13
S15         76    S10 OR S11 OR S12 OR S14
S16    3894081    TEXT OR TEXTUAL OR DOCUMENT? OR WORD? ? OR PHRASE? OR PARA-
                  GRAPH? OR TERM OR TERMS OR (CHARACTER OR TEXT)(N)(STRING? OR -
                  SEGMENT?)
S17      68276    S3(2N)S16
S18      26648    S16(2N)(MATCHING OR SIMILAR OR SAME OR IDENTICAL OR EQUAL -
                  OR EQUIVALENT)
S19          1    S18 AND S4 AND S5 AND (S6 OR WORKSPACE? OR WORK()(SPACE? OR
                  AREA?) OR CLIPBOARD?)
S20          1    S18 AND S4 AND (S6 OR WORKSPACE? OR WORK()(SPACE? OR AREA?)
                  OR CLIPBOARD?)
S21         47    S18 AND S5 AND (S6 OR WORKSPACE? OR WORK()(SPACE? OR AREA?)
                  OR CLIPBOARD?)
S22        291    S18  AND (S6 OR WORKSPACE? OR WORK()(SPACE? OR AREA?) OR C-
                  LIPBOARD?)
S23         59    S22 AND (S7 OR S8 OR S9)
S24        171    S15 OR S19 OR S20 OR S21 OR S23
S25        132    RD (unique items)
S26         88    S25 NOT PY>2001
S27         12    S15 AND S17
S28          9    S27 NOT S26
S29          5    RD (unique items)
S30          1    S29 NOT PY>2001
S31         10    S18 AND JOIN? AND S4
S32         10    S31 NOT S25
S33          7    RD (unique items)
File    8:Ei Compendex(R) 1970-2005/Jun W1
           (c) 2005 Elsevier Eng.  Info. Inc.
File   35:Dissertation Abs Online 1861-2005/May
           (c) 2005 ProQuest Info&Learning
File   65:Inside Conferences 1993-2005/Jun W2
           (c) 2005 BLDSC all rts. reserv.
File    2:INSPEC 1969-2005/Jun W1
           (c) 2005 Institution of Electrical Engineers
File   94:JICST-EPlus 1985-2005/Apr W4
           (c)2005 Japan Science and Tech Corp(JST)
File  111:TGG Natl.Newspaper Index(SM) 1979-2005/Jun 14
           (c) 2005 The Gale Group
File    6:NTIS 1964-2005/Jun W1
           (c) 2005 NTIS, Intl Cpyrght All Rights Res
File  144:Pascal 1973-2005/Jun W1
           (c) 2005 INIST/CNRS
File   34:SciSearch(R) Cited Ref Sci 1990-2005/Jun W1
           (c) 2005 Inst for Sci Info
File   99:Wilson Appl. Sci & Tech Abs 1983-2005/May
           (c) 2005 The HW Wilson Co.
```

5256047   INSPEC Abstract Number: C9606-5260B-164
  Title: Evaluating spatial correspondence of zones in document recognition systems
  Author(s): Garris, M.D.

  Author Affiliation: Nat. Inst. of Stand. & Technol., Gaithersburg, MD, USA

  Conference Title: Proceedings. International Conference on Image Processing (Cat. No.95CB35819)   Part vol.3   p.304-7 vol.3

  Conference Title: Proceedings International Conference on Image Processing
  Conference Sponsor: IEEE Signal Process. Soc
  Conference Date: 23-26 Oct. 1995   Conference Location: Washington, DC, USA
  Language: English   Document Type: Conference Paper (PA)
  Treatment: Theoretical (T); Experimental (X)

  Abstract: This paper introduces scoring methods developed to automatically assess the performance of document recognition systems, specifically, to evaluate the spatial correspondence of zones produced by a **document segmentor** . **Two** different approaches are discussed. The first approach (based on zone overlap and nearest-neighbors) is better applied to **merged** zones, whereas the second approach (based on zone alignments) is better applied to nested zones (such as those found in tables and graphs). Definitions of coverage and efficiency error are presented, and scoring results on real system output is provided that validates the usefulness of these methods to **compare** different **document** recognition algorithms. Currently, no standard testing procedures exist for measuring and **comparing** algorithms within a complex document recognition system. Scoring methods, like the ones introduced in this paper, serve as design and validations tools, expediting the development and deployment of document analysis technology for system developers and end users. (13 Refs)
  Subfile: C
  Descriptors: document image processing; image recognition; image **segmentation**
  Identifiers: spatial correspondence; document recognition systems; document **segmentor** ; zone overlap; nearest neighbors; **merged** zones; zone alignments; nested zones; tables; graphs; coverage; efficiency error; scoring results; system output; document recognition algorithms; scoring methods; validations tools; design tools; document analysis technology
  Class Codes: C5260B (Computer vision and image processing techniques); C6130D (Document processing techniques)